

News Analysis

Internet of the future may be a one-stop information shop

Data sellers such as Dow Jones are setting up shop on the world's largest data network.

By Margie Wylie

Whether it's the wind-speed velocity of an unladen African swallow or the asking price for pork bellies, nearly anything you want to know today is offered for sale on-line somewhere by someone. Finding out who has that information and negotiating the connection, however, is a different story. With literally hundreds of vendors hawking millions of pieces of information on myriad public and private commercial networks, there exists no comprehensive, uniform and simple way for users to search out the information they need or connect to the specific service where it resides.

That may soon change. As the Internet allows more commercial users on-line, data peddlers such as Dow Jones & Co. Inc. of Princeton, N.J.; Mead Data Central of Miamisburg, Ohio; and Dialog Information Services Inc. of Palo Alto, Calif., are beginning to set up shop on the worldwide network that was once the exclusive territory of military, research and education users.

Should a new Internet standard for searching databases emerge, Mac users could soon find themselves only a command away from solving whatever puzzles them.

But don't apply for a node just to do database searches, yet. Most information providers still are offering only limited Internet access, and they don't seem to be in a hurry to expand.

The way it is. Today, most big information sellers require users to either connect to their mainframes directly, through a private network or via proprietary and sometimes expensive networks such as CompuServe, or by dialing into Tymnet or other commercial networks.

Users can connect to only one service at a time, they usually have to learn a proprietary and often cryptic interface for searching each service, and they usually are charged even when a search isn't successful.

The Internet, through such search services as Wide Area Information Server (WAIS), offers users the prospect of searching several different providers without logging onto different networks and using the same standard, albeit cryptic, commands they use to get at any other Internet data source. Network services such as Telnet let users link up to commercial databases through the systems' native interfaces.

With an increasing number of

Internet-connected networks willing to carry commercial traffic, it has become a viable alternative for corporations to connect to a variety of on-line database services.

"The Internet has sort of snuck up on the commercial world," said Taylor Walsh, an independent consultant based in Washington, D.C., and author of *New Commerce*, a report on emerging business opportunities on the Internet.

Today many big on-line information providers are struggling to meet the commercial users' demand for Internet access, Walsh said.

According to Walsh's estimates, there are about 3 million commercial Internet users in the United States, and the network is still growing, as it has for the past two years, at a rate of more than 20 percent per month. (Other estimates of the growth rate range from 10 percent per month to 25 percent.) With a push for liberalization of Internet usage policies, which have been particularly hostile to free commerce, "the majority of organizations getting connected in the last 18 months have been commercial organizations," Walsh said.

Today, three of the biggest on-line information providers, Dow Jones, Mead Data Central and Dialog, offer some level of Internet access to some users. Dow Jones offers only its News and Retrieval Service, which consists of about 70 databases, via Internet connections. Most of its customers are still academics and researchers, the company said. However, Dow Jones said the interest from commercial users is soaring.

Mead Data Central claims it offers access to its Lexis law database over the Internet only to law schools. But it reportedly is beginning to sell Internet access to both its newspaper database, Nexis, as well as Lexis, to some commercial users who can log onto the services via Telnet.

Both Mead Data Central and Dow Jones offer their services via JvNCnet (John Van Neuman Computer Network), a Princeton, Mass.-based network that charges users for access time in addition to expensive database charges. However, the companies already are realizing a lower operating cost by doing so. Access through commercial Internet providers is more economical than other access methods, said Maggie Landis, spokeswoman for Dow Jones.

Dialog, a Knight-Ridder information service, offers its services through ANSnet, a partially government-funded network that by nature doesn't yet allow unrestricted commercial traffic. Dialog supports only academic and research users on ANSnet, for now.

The way it could be. Perhaps the brightest promise for Internet access, however, lies with a new technology developed jointly by Apple, Dow Jones and Thinking Machines Corp. WAIS is a full-text data-retrieval standard that will let users search for the information they need over the Internet (see MacWEEK, Oct. 26, 1992).

Using a single English-like query, users can search for what they need among about 340 WAIS servers on the Internet, then pay only for what they use. Users don't have to navigate the Internet to get to a particular network, such as JvNCnet, and then log onto a service through Telnet. They just send out a request scattershot-style and wait for replies from the WAIS services.

Currently, there are more than 200 free databases available via WAIS over the Internet, according to Seybold Publications Inc.'s Digital Media newsletter. Even the Library of Congress plans to take its massive catalog on-line through WAIS. With a WAIS Mac interface already available, users can begin data-surfing WAIS servers today.

The way vendors see it. Yet despite the promise of Internet access and WAIS, many providers appear to be dragging their feet in making more than academic and research, or very limited commercial access, available. WAIS services are considered blue-sky investments by these companies.

Logistically, setting up Internet access takes an investment in time and money that these companies want to be certain will pay off, said Jim Joseph, a spokesman for Mead

Data Central. Many providers have to set up TCP/IP-based access to aging and proprietary mainframe programs designed for X.25 packet-switched connections and proprietary search protocols.

With WAIS servers, vendors have to redefine completely how they charge for their services and settle issues of copyright and usage that arise when their databases are so widely available, Walsh said. But he said he thinks that commercial users could benefit from the work already under way on these problems in public forums. Unlike free databases, there currently is no WAIS server code in the public domain for profit-driven information servers.

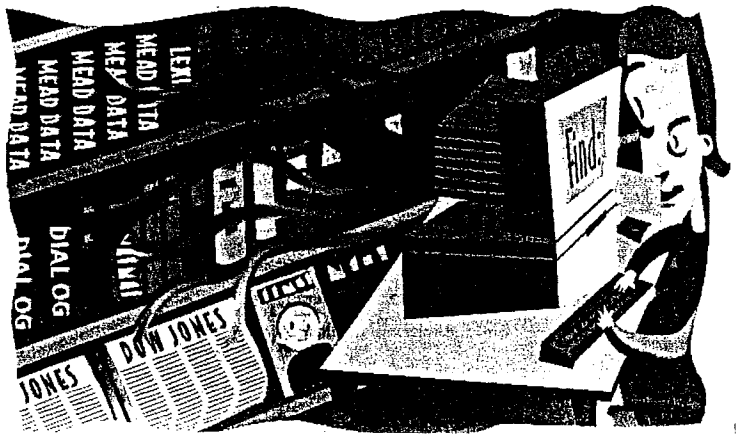
There also seems to be fear among commercial information providers that the Internet may yet collapse under the weight of its own success.

"From our perspective, one stumbling block is the growth of the Internet," said Bill Frambes, manager of the technical-support group for Mead Data Central.

The network is growing so quickly that it will soon run out of addresses, according to Jeffrey Schiller, network manager for the Massachusetts Institute of Technology in Cambridge and an active member of the Internet Engineering Task Force, a collection of committees that sets the standards for using the TCP/IP network. The IETF is currently working to extend the Internet's addressing scheme, but "D-day is measured in months not years," Schiller said.

Vendors are skittish about putting their chips on a network that has no central control and whose policies can change on a political whim, company representatives said. Although the network has indeed become more open to commercial traffic, the original NSFnet backbone and now ANSnet, which replaced it, is still subject to a usage policy that prohibits commercial traffic from even passing through on its way elsewhere.

See Internet, Page 24



Internet

From Page 22

Advanced Network Services Inc., an IBM Corp. and MCI Communications Corp. joint venture with the nonprofit Merit Networks Inc., was charged with opening the backbone to a wider range of traffic when it took over management of the Internet in 1989, but so far it has been locked in bitter disputes with other networks about how much it can charge other providers for use of its backbone.

Nevertheless, users can't help but suspect that there's something else behind the reluctance of commercial database providers to play on the giant, nearly free Internet.

Are these giants simply putting off the day when they have to compete head to head in an open information market? Providers today charge connect time, search time and download fees that can climb easily into the hundreds of dollars even for a relatively fruitless inquiry. Under WAIS, they might be passed

over for a service that provides the right information at a better price.

Information service representatives denied the charge: "We're not in the telecom business; we're in the information business. We want people to have as many channels as possible," said Libby Trudell, director of product marketing for Dialog.

In any case, most information providers agree that commercial database access isn't a matter of "if" but "when." □

Internet and some innovative software are changing the nature of the library. Instead of being at the end of the block, a book collection will exist in cyberspace.

Good-bye, Dewey decimals

By David C. Churbuck

IT WOULD BE TOO much to argue that the jobs of the 152,000 librarians in the U.S. are in jeopardy. But it's fair to say that their jobs will change dramatically over the next two decades, courtesy of the Internet computer networking system and new software that controls Wide Area Information Servers, or WAIS.

Internet, a global network of computers formed by the government to connect universities, research labs and military complexes, was until recently largely limited to nonprofit users. But it is now taking on commercial customers (*FORBES*, July 8, 1991), who are finding the system an extremely economical way to gather and trade information. For as little as \$1 an hour, a subscriber to Internet can sit at a personal computer in his office, issue a request for information, and have the network route his request to libraries across the globe. The system will retrieve in a matter of seconds a collection of card-catalog citations that would have taken a lot of shoe leather to find in person. How else would a scholar in Chicago find out that "chaos" was touched on in a book written in 1741 in Latin and found in Harvard's library? Or what's in the Australian Defence Force Academy library. If he wants the volume, the researcher can follow up with a request for an interlibrary loan.

Full-text retrieval, still very limited at this point, is around the corner. When it comes, the local library as we know it all but disappears. In lieu of librarians we will have programmers and database experts.

One of the fathers of the Wide Area Information Servers concept is Brew-



Acy Harper

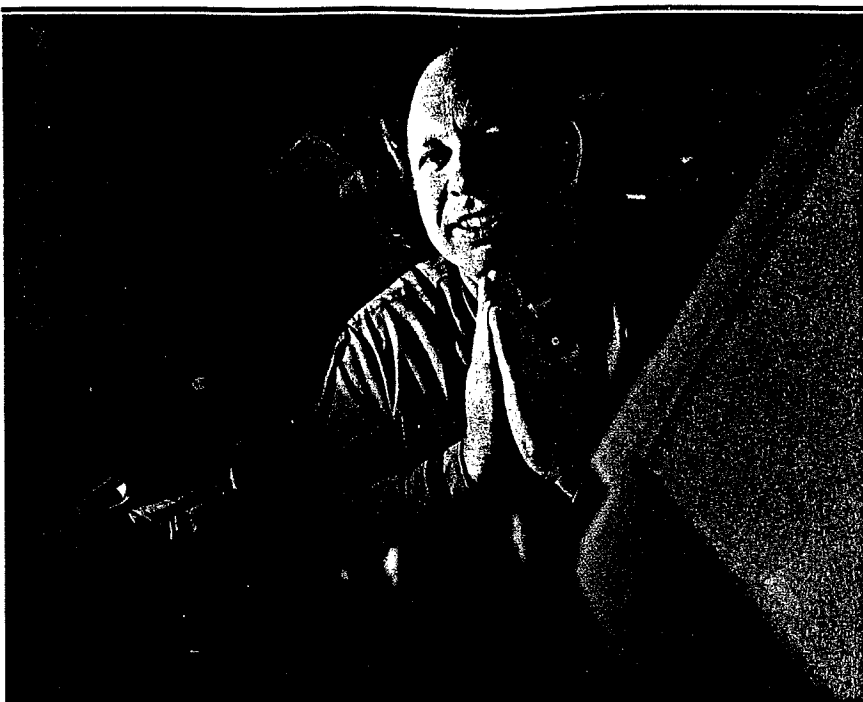
WAIS Inc. founder
Brewster Kahle
**Building the
electronic library
of the future.**

ster Kahle, who developed programs for searching databases when he worked at Thinking Machines Corp., the Cambridge, Mass. manufacturer of supercomputers. Thinking Machines is primarily interested in developing a market for its hardware, whose parallel processors are ideally suited to massive text searches. Kahle's interest is in the software.

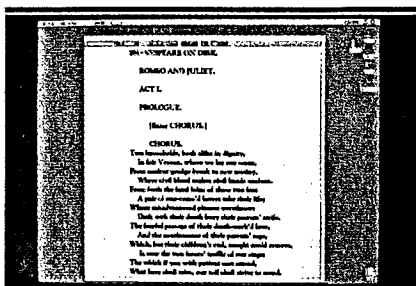
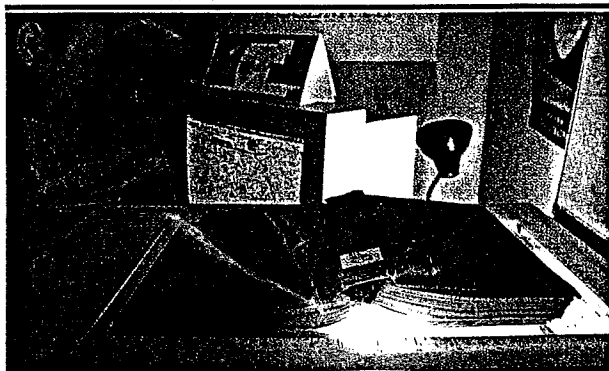
"People want to be able to pose a

question to the net[work] and not care where the answer comes from," says Kahle, 33, who recently left Thinking Machines to form WAIS Inc., in Menlo Park, Calif. Kahle's firm helps corporations and publishers set up electronic versions of copyrighted and public domain material and then make it accessible to users who are willing to pay for each download.

Kahle's software first paws through



Project Gutenberg founder Michael Hart; book scanning
Filling the shelves.



the raw text, indexing it to speed up later searches. Users can then submit plain-English queries. Ask for songs with the words "eyeball" and "toe jam," and it will come up with the Beatles' "Come Together." The system learns from trial and error, since the user can tell the computer which of the retrieved documents was closest to what he was looking for.

"WAIS is like a research librarian

who watches you read through a stack of information, taking notes on what you looked at first and set aside for future reference, and what information you threw away," explains Kahle. You can use WAIS to search a database on your own computer, but the software is optimized to handle the complexities of searches through physically distant databases.

Programmers at the University of Minnesota have taken WAIS a step further. To get information from a WAIS, you have to know exactly where that information is—which server has, say, a collection of song lyrics. The new software, called Gopher, makes it easier to navigate through the network. Richard Wiggins, Gopher coordinator at Michigan State University in East Lansing, Mich., explains the difference between WAIS and Gopher:

"WAIS is for direct searches. Gopher is a browsing tool." Enhancements to Gopher aim at Kahle's ideal—a query about the Beatles is answered with simultaneous searches through all the card catalogs on the network.

More libraries are joining the two-year-old Gopher system all the time, but the best is yet to come, including the 22-million-volume Library of Congress.

As with any new technology, some librarians are finding it hard to adapt. Says Wiggins: "In general terms, this technology is of interest to libraries, but it is frightening at the same time, setting off a struggle between academics who want knowledge spread around and librarians who want to control it."

Wide area information fanatics dream of an interlibrary loan system on a grand scale, in which only one copy of each book or magazine is in the system at all. That one copy could be used by many people at the same time and it would never be lost or overdue. Also, to the consternation of publishers thinking about royalties, it could be easily duplicated.

For this system of the future to become real, existing libraries of printed books will have to be digitized. That's the objective of Project Gutenberg, the creation of Michael Hart, professor of electronic text at Illinois Benedictine College in Lisle, Ill. If Kahle is building the shelves for the electronic library of the future, Hart is filling them. Gutenberg transcribes books, mostly in the public domain, for free digital distribution, whether via floppies or Internet.

For most of the 22 years that Hart has been at work on Gutenberg, digitizing has meant the slow and arduous retyping of the classics. Now, however, all Hart has to do is shred the binding off a book and feed the sheets into a scanner. Gutenberg has 50 titles on-line so far and expects to have 10,000 ready by the year 2001. That's nothing next to what the Library of Congress could crank out if it got serious about digitizing its collection, a project just getting under way. And if the programmers can solve some knotty problems involving pay-per-view and copying, there's no reason that newly published books cannot go on the system, too.